Explorations in a Classroom Context of 3D Audio for Game Environments

Greg Corness* — Columbia College Chicago
Kristin Carlson † — Illinois State University
Zachary Buckley ‡ — Illinois State University

Abstract

This work explores how current environments for game programming are supporting a wide variety of audio options. We question whether existing approaches such as HRTF are giving any more spatial information than we are getting from a distributed virtual system. We compare current theory in physical experiences of spatialized audio with existing options in game environments to analyze how realistically we can gather the physical information used to experience audio in 3D.

Index Terms: Applied computing-Sound and music computing—Applied computing Performing arts—Arts and Humanities Sound and music computing—;

1 Introduction

Audio in media environments has developed in parallel but somewhat separately from visual techniques. If we consider the developmental progression of sound in media, VR, and games we can break it down into: 1) sound association through synchronous events, 2) associated sounds in the virtual space (assets mapped to basic mixing and specialization), and 3) 3D processing on sound assets. As the tools for creating virtual environments have been developing, so has the desire to include sound and construct both artistic and technical techniques for including sound as an integral part of the immersive experience.

This paper presents our experience of designing audio environments based on three projects in a classroom environment. All three projects were conceived as explorations in audio-only video gaming. The projects were realized in Unity and explored the built-in spatializers and 3D audio plugins as well as third party plugins for 3D audio.

2 Background

In the acoustic world our two ears enable us to separate and locate the sounds around us. This ability is a normal part of our experience of our world. Our auditory perception compares the sound sensed in each ear by processing differences in parameters such as volume, timbre, and time (phase). When perceiving the sound in the environment this information allows the listener to place the sound in 3D space; including the height of the source, the distance to the source and its position around us (in front or behind) [3].

Media sound environments have attempted to provide an illusion of the acoustic experience [1]. This work started with the development of stereo and quadraphonic recordings. In general this work can be seen as attempting to record or add space to a media sound experience often referred to as spatializing the sound. In more recent years this work has pushed forward and to what is referred to as 3D audio [2].

In general we can understand the difference in approaches in this way. Spatialized sound uses digital signal processes to add the illusion of space to a audio signal [6]. This could include reverb attenuation (volume) or other filters. In contrast, 3D audio attempts to consider the difference in the perception of the two ears for each sound. This includes aspects such as the timbral shadow (reduced high frequencies) that the far ear is subject to as a result of the head blocking high frequencies, the reduced volume of the audio in the far ear as a result of its increased distance from the sound source in relation to the near ear, and the delayed time to hearing the signal in the far ear [5]. The most common approaches currently in use are ambisonic and HRTF.

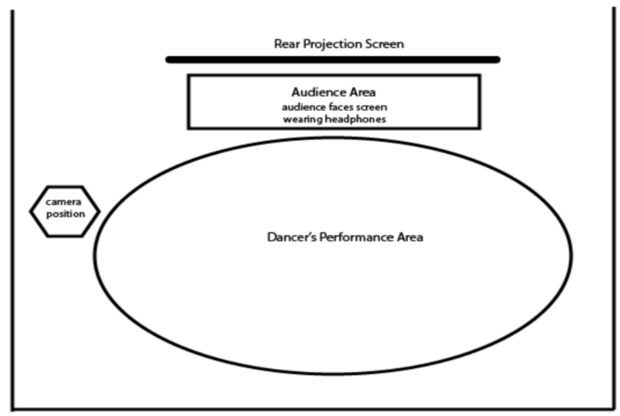

Figure 1: A depiction of the stage setup for the Illusion of Togetherness project. The image is from [4] and we have permission to reuse it.

In the paper The Illusion of Togetherness the authors show a perceptual connection between audiences of mediated performance [4]. The study was setup with the participants facing a large screen, with a camera offset of the performance taking place behind them (see Figure 1). The participants wore headphones, so that both their visuals on the screen and the audio of the performer behind them could be mediated through delay, overlap, and synchronicity. The research suggested that senses that are perceived at the same time are believed by the participant as representing the reality of the situation prompting other sensations to be understood as fabricated by the situation. For example, if an audience member hears the sound of and actor jumping then sees and feels a mediated version of the actors actions, they reported believing that the audio was faked and the visual and kinetic information reliable as indicating when the actor jumped. This research indicates the strong reliance on synchronization when we cognitively build our perception of and experience. This work was performed in a mixed reality situation but the it is reasonable that the results hold for a virtual environment.

3 Experiments

We explored three experimental projects in a game class environment to explore the experience of sound through a variety of audio designs. The class had nine students, eight male and one female. All 9 students were sound for game majors with experience implementing sound in unity but had not used the 3rd party 3D audio plugins before. The first project (Audio-Only Navigation Game) was made

*e-mail: gcorness@colum.edu
†e-mail: kacarl1@ilstu.edu
‡e-mail: zwbuckl@ilstu.edu

by one student and critiqued by the entire class. The second project (Basic Listening Project) was created by the instructor and critiqued by the entire class. The whole class worked together to create the last project (Audio-Only Avoidance Game) and critiqued it as a class.

3.1 Audio-Only Navigation Game Project

The first project involved the building a multilevel game that relied on audio for all aspects of the game play and of special interest the navigation of the space. The sounds in the game were all derived from recordings of a violin including various extended techniques. This sound pallet was used as a unifying element to the design but was not tested for clarity. The exploration of designing with 3D audio focused on how the audio could provide information required for navigating a complex space. Sound objects such as machine, computer terminal, doors, ventilation noises were placed in the environment to provide sonic landmarks for navigation. Sounds were also created for different walking surfaces including surfaces to jump from and onto. Such design elements are standard in game sound design. What we were more interested in is if playing these elements through an HRTF 3D sound engine would improve the experience for the player 1) to provide a deeper sense of immersion, 2) in providing a clearer sense of direction by being able to sense the sound in close to 360 degrees. Based on initial tests, a third party HRTF plugin to Unity was used.

At the conclusion of this exploratory project we found that our initial skeleton design of a few basic sounds had promising responses from users. However, as we added sounds we found the navigation elements became less distinct especially those requiring perception of height. We also found that technical difficulties seemed to mound. It was difficult to tell if the sounds and spatialization was still working. The mix of sound elements became unstable and occasionally would not play. From a designers perspective we found the affordance of the tool though promising in the beginning, in the end not worth the effort.However, we remain interested in what issues were at play and if there are design techniques for working around the current limitations of the tool.

3.2 Basic Listening Project

In order to try to isolate some of the advantages and disadvantages discovered in the Navigation project, a simple project was undertaken as a proof of concept and test. This base project attempted to isolate some of the issues discovered in the Navigation project by exploring the basic perception of directionality afforded by using a 3rd party 3D audio plugin employing HRTF (RealSpace 3D Audio). In this project a sound environment was constructed with three sound objects: a static object at floor level, a static object above the players head (but located at a point), and a third wandering sound object. Users were asked to try to point to the different sound objects and to capture the moving object. Again this game was presented as a audio only game although the design did have visuals that could be reference for comparison. Although the software allowed for customizing the HRTF settings these were used to the best of the user's ability.

The results of this project supported the initial experience, that the 3D audio did provide a heightened experience of sound emanating from behind the user. This behind the back sound was experienced with both the stationary and moving sounds. However, the sound object that was raised above the floor was found to be much harder to place. We also found that though the user could turn and experience the sound going behind, it was much more rare for the user to experience the moving object as coming up behind them. The reported experiences suggest that two issues may be at play. 1) the tracking of a sound from the side to the back space enforces the perception of the sound coming from behind, and 2) that a limited perception of depth means that the experience is the sound appearing behind rather than sneaking up behind. Based on these results we became interested in exploring other design approaches and uses.

3.3 Audio-Only Avoidance Game Project

The third project focused on how 3D audio could provide information on the location and movement of an object in the Virtual space focusing on depth and proximity. Of particular interest was the distance away from the user that the object was moving. The project was realized using Unity and was designed around a blind frogger concept. The object of the game was to cross a busy road without the aid of visual cues only using the audio cues. The game included a training level and two play levels.

Based on our earlier projects we tested not only third party 3D sound plugins but looked again at spatialization and built in 3D sound plugins for Unity. Our initial design again focused on using the plugin RealSpace 3D Audio but half way through development the plugin started giving us problems so we switched to the build in Oculus spatalizer. We also explored alternative design techniques to address issues found to be weak when addressed using only the plug ins.

The game design of Blind Frogger enabled us to focus on proximity as the primary test element. As with any sound design UX sounds such as leveling up, win state, death, and re-spawning were all designed to be clear. Though these were also spatialized there was no comment on the experience changing with these sounds.

The main consideration for the audio design for this game was providing the user with the information of how close a car was. We found that the 3D audio provided a nice experience of space that suggested distance. However, the experience of distance did not translate into information that made it possible to avoid the cars. The sound team found that the proximity to the car was not strong enough to help avoid a collision and so devised a method for providing a sound when the car was heading at the user different from passing the use. With this design shift it was then found that simple spatialization was sufficient to make the game mechanic work and simplified the technical difficulties of using the 3D audio.

4 Conclusion

In this paper we have presented our work with three projects exploring the experience reliability and effectiveness of current sound engines for providing spatial, locative, and navigational information to users in game style VR environments. We have focused on audio only environments as a method for critically assessing the tools and the design concepts being explored. These projects were completed largely as pedagogical and exploratory tasks. There is much more work to be done to fully explore the ideas we have presented. Furthermore, third party tools for 3D audio in games and VR environments are under rapid development and many of the techniques are still being researched. Based on our experiences developing the presented projects we suggest that audio-only games and environments provide a uniquely critical assessment of the tools. We found that visual cues were very powerful in convincing the user of the correctness of the virtual positioning of the audio [4]. However, in the audio only tests the virtual position was found to be less precise and subject to confusion in complex environments. We also found that by designing sounds with such limitations in mind strengthened the overall effectiveness of the design, but these techniques were of similar effect for both spatialized and 3D sound tools. Still, by continuing to work at creating audio only environments using VR and game design tools, we can provide critical assessment and help guide the continued development of tools and design techniques for audio in VR environments.

References

[1] D. R. T. Begault. 3-D Sound for Virtual Reality and Multimedia. Technical report, Aug. 2000.

[2] J. Breebaart and C. Faller. *Spatial Audio Processing: MPEG Surround and Other Applications*. John Wiley & Sons, Mar. 2008. Google-Books-ID: vxP4q0RE__oC.

[3] A. S. Bregman. *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press, 1994. Google-Books-ID: jI8muSpAC5AC.

[4] G. Corness, K. Carlson, and T. Schiphorst. Audience Empathy: A Phenomenological Method for Mediated Performance. In *Proceedings of the 8th ACM Conference on Creativity and Cognition*, C&C '11, pp. 127–136. ACM, New York, NY, USA, 2011. doi: 10.1145/2069618.2069641

[5] J.-M. Jot. Real-time spatial processing of sounds for music, multimedia and interactive human-computer interfaces. *Multimedia Systems*, 7(1):55–69, Jan. 1999. doi: 10.1007/s005300050111

[6] G. S. Kendall. Spatial Perception and Cognition in Multichannel Audio for Electroacoustic Music. *Organised Sound*, 15(3):228–238, Dec. 2010. doi: 10.1017/S1355771810000336